

<https://helda.helsinki.fi>

---

## The ethics of behavioural public policy

Lepenies, Robert

Routledge  
2018

---

Lepenies , R & Malecka , M 2018 , The ethics of behavioural public policy . in A Lever & A Poama (eds) , The Routledge handbook of ethics and public policy . Routledge handbooks in applied ethics , Routledge , Abingdon , pp. 513-525 .

---

<http://hdl.handle.net/10138/308706>

---

acceptedVersion

---

*Downloaded from Helda, University of Helsinki institutional repository.*

*This is an electronic reprint of the original article.*

*This reprint may differ from the original in pagination and typographic detail.*

*Please cite the original version.*

***Abstract.*** Public policy is increasingly informed by insights from the behavioural sciences. We highlight three aspects of behavioural public policy which can be incompatible with democratic ethics. First, policy-makers can use behavioural instruments such as nudges to steer citizens' behaviour without giving reasons and by relying on non-participatory research methods. Second, behavioural public policy is frequently implemented in the form of administrative discretion by organisations which are not under direct democratic control. Third, behavioural public policy currently relies on a partial and narrow view of the behavioural sciences and is inattentive about value judgements already embedded in the research it draws upon.

**It is an almost final version of the paper published as: Lepenies R., Malecka M. (2019), Ethics of behavioural public policy, [in:] Poama A., Lever A., Routledge Handbook of Ethics and Public Policy, Routledge, pp. 513-525**

Running Head Right-hand: The ethics of behavioural public policy

Running Head Left-hand: Robert Lepenies and Magdalena Malecka

## 40

# THE ETHICS OF BEHAVIOURAL PUBLIC POLICY

*Robert Lepenies and Magdalena Malecka*

### **What is behavioural public policy?**

Insights from the behavioural sciences, such as cognitive psychology and behavioural economics, are currently reshaping much of public policy around the world. Behavioural approaches are drawn upon in a variety of policy fields, such as in health and environmental policy, labour regulations, and consumer protection law. The first nudge unit, the Behavioural Insights Team (BIT), was established in 2010 under David Cameron's coalition government in the UK within the cabinet office. It has since become a quasi-privatised company ("Behavioural Insights Limited") with more than 150 employees that consults and tests policy innovations for the UK government and abroad. The BIT serves as a model for many other behavioural insights teams worldwide: in the US, Germany, France, the Netherlands, Australia, Japan, and Singapore, which all have recently created nudge units. Similar behavioural policy units can be found at the World Bank and at different teams within the United Nations, at the OECD, and the European Commission (Joint Research Centre). A peak point in behavioural sciences applied to policy was reached when President Obama issued, in 2016, an executive order entitled "Behavioral Science Insights Policy Directive" (E.O. 13707, 2016) that aimed at developing 'strategies for applying behavioural science insights to programs and, where possible, [to] rigorously test and evaluate the impact of these insights'. Shortly before, the White House had created a Social and Behavioral Sciences Team (SBST), which was influenced by the tenure of Cass Sunstein as head of the Office of Information and Regulatory Affairs during the Obama administration. The take-up of psychological insights by public actors is noteworthy as it often presents a far-reaching change from the status quo of the contemporary design, formulation, implementation, and evaluation of policy (Halpern, 2016).

The first normative defence of behavioural public policy (BPP) was spelled out in the writings on libertarian paternalism (Thaler & Sunstein 2004) and in *Nudge* (Thaler & Sunstein 2008). Ethicists of public policy have, for now more than a decade, meticulously engaged with the framework formulated by Thaler and Sunstein, creating a large and sprawling literature, which we would like to divide into “general philosophical”, “topical”, and “discipline-specific” bodies of work. Most *general* philosophical treatments can be found within moral and political philosophy (Hausman & Welch 2010; Wilkinson 2013). For instance, the defensibility of libertarian paternalism has been called into question (Mitchell 2004; Rebonato 2013), while conceptual debates have sought to define nudging (Saghai 2013; Hansen 2016) and related notions such as choice architecture (Vallier 2014). At the same time, distinct *topical* discussions about the ethics of nudging have emerged independently from within fields in which nudges have been applied. Examples include debates in sub-fields such as public health (Menard 2010) and bioethics (e.g., in clinical contexts when discussing nudging by physicians, see Cohen 2013; Gorin et al. 2017), ethics of artificial intelligence and big data (Helbing et al. 2017; Yeung 2017), environmental nudges (Schubert 2017b), and charitable giving (Hobbs 2017) and development (Bernard 2014). Here, debates are specific to the field of application and rarely discuss the ethics of nudging across fields in general terms.

In addition to general and topical debates about the ethics of nudging, we find reflections about the theoretical and methodological compatibility of behavioural approaches for a given scientific discipline. Examples include debates in law and legal theory (Alemanno & Sibony 2013; Kemmerer et al. 2014), economic methodology and welfare economic theory (Sugden 2017; Whitman & Rizzo 2013), development studies (Reddy 2013), health policy (Dingley 2013), cognitive and social psychology (Gigerenzer 2015; Hertwig & Grüne-Yanoff 2017), and marketing (French 2011; Christ 2014). The discipline-specific body of literature about the ethics of nudging shows that not just policy but also the social sciences are being ‘behaviouralised’ as experimental methodology and behavioural insights are spreading throughout academia (Malecka & Lepenies 2018).

Yet, the contributions mentioned earlier mostly take Thaler and Sunstein’s framework and concepts not just as the starting point but also as the end point of their engagement. The current debate on ethics of nudging is focused on defending or criticising this initial theoretical framework.

We here want to argue that ethicists of public policy must look elsewhere in the future: they should ask how public policy is actually undertaken – that is, how behavioural policy is institutionalised and argued for in practice, and – being aware of its intellectual origins and context – begin to reflect on the justification of behavioural interventions if these are to become more widespread. It is helpful in this regard for ethicists of public policy to take note of incipient research *on* BPP (whether in the sociology of BPP or in the epistemology and intellectual history of the behavioural sciences). We will focus on three aspects of BPP – namely, the instruments it uses, the organisational forms it involves, and the view on (behavioural) science it presumes. BPP poses different problems in democracies: it is hard to make it compatible with deliberative and participatory approaches in research and implementation, it often relies on instruments that impact the behaviour of citizens without necessarily relying on reason-giving, and

it instrumentalises science for practical use, potentially eroding public trust in science and scientists.

The chapter is structured as follows: We begin by contrasting behavioural public policy with evidence-based policy-making, describing the former as an increasingly institutionalised policy movement. Then, we describe how behavioural public policy invokes “science” while justifying behavioural interventions. We then consider the challenges to democratic ethics posed by behavioural policy instruments, behavioural organisations, and its relation to science before concluding.

### **Behavioural public policy and evidence-based policymaking**

The phenomenon of behavioural public policy is closely related, though not identical, with that of evidence-based policymaking (EBPM), which is an approach to policy broadly encouraging a principled grounding of its decisions on the best available scientific evidence. BPP, like EBPM, aims to ground policy design, its formulation, implementation, and evaluation on scientific insights. BPP restricts these to scientific insights about human behaviour won from the “behavioural sciences”, whereas EBPM is not limited to scientific insights stemming from a specific domain or set of scholarly approaches. Both BPP and EBPM accept the ‘specific hierarchy of scientific methods, with randomised control trials (RCTs) and meta-analysis/the systematic review of RCTs (published in high-status peer-reviewed journals) at the top’ (Cairney 2016: 3). In practice, this is expressed particularly in the endorsement of experimental methodology (especially through RCTs) and meta-studies.<sup>1</sup>

What matters to us here is that EBPM is most commonly described as a ‘vague, aspirational term, rather than a good description of the policy process’ (Cairney 2016: 1). This distinguishes BPP, which represents both a novel intellectual approach to policy and a policy movement with increasingly well-defined contours. For a good example of this behavioural proposition, see the call for bridging the ‘divide between behavioral science & policy’ in Fox and Sorkin (2014). Proponents of BPP aim to change the way in which academic scholars influence policy, and understand themselves as a ‘growing movement among social scientists and leaders within the public and private sector, dedicated to grounding important decisions in strong scientific evidence’ (Behavioral Science & Policy Association 2017).

BPP is increasingly institutionalised, networked, and embedded in governance practices: it has become a policy movement. Studies of networks of behavioural change agents show the spread of a transnational network of BPP proponents with a distinct rhetoric, journals, and associations alongside and beyond new governmental “nudge units” that are created around the world (Jones et al. 2013; Rykoff et al. 2016; Strassheim & Korinek 2016; Whitehead et al. 2014). With varying zeal, entrepreneurial proponents of behavioural policy push towards bringing behavioural insights to policy (John 2014), making sophisticated calculations comparing the relative effectiveness of behavioural instruments to alternative policy measures (Benartzi et al. 2013).<sup>2</sup> While doing so, traditional (non-behavioural) approaches to policy are commonly labelled as not being rigorous enough. Hence, it has become a frequent strategy for behavioural insights teams to quantify and monetise their achievements in terms of how the application of the behavioural scientific findings leads to tax dollars saved, pollution avoided, accidents prevented, lives saved, and happiness increased. Indeed, using behavioural

science is defended as a moral imperative. As one of the leading behavioural analysts of the French prime minister's Centre for Strategic Analysis argued, 'No one would accept that a new drug would be developed only by economists and lawyers and launched without the proper trials. We should not tolerate this in policy-making either' (Oullier 2013: 463). Because of these normative assumptions of BPP, ethical reflection is needed.

### **Intellectual underpinnings of behavioural public policy**

All key behavioural publication outlets are keen on discussing the ethics of behavioural policy, with frequent invitations to ethicists and applied moral philosophers to comment (in addition to frequent inclusion of such perspectives at behavioural conferences). On the key tenet, however, proponents of behavioural public policy are not swayed: the consensus is that choice architecture is inevitable; there is no way not to nudge. For example, Sunstein argues,

It is pointless to raise ethical objections to nudges and choice architecture. [. . .] No government can avoid some kind of choice architecture. We can object to particular nudges, and particular goals of particular choice architects, but not to nudging in general [. . .] government is nudging even if it does not want to do so.

(Sunstein 2014: 15–16)

Indeed, in the “age of the behavioral sciences” (the title of Sunstein 2016) it would be irresponsible not to use science for the greater societal good.

Yet, little has been said about what it means that policy is informed by the “behavioural sciences”. What counts as a behavioural science in this movement? BPP is drawing on a specific subset of disciplines that has historically included some but not other strands of understanding human behaviour (Heukelom 2014; Thaler 2015). BPP is consciously interdisciplinary, but selectively so (Lepenies & Malecka forthcoming). Put a bit provocatively, BPP itself has a selective science bias. Only certain approaches are being received that share the methods and styles of inquiry of BPP: those that heavily draw on behavioural economics (which itself developed out of cognitive psychology – e.g., Tversky & Kahneman 1975 or Kahneman & Tversky 1979) and use experimental methodology (see Malecka & Nagatsu forthcoming on the variety of approaches within the behavioural sciences). BPP is almost entirely influenced by a specific strand of study of human behaviour (cognitive psychology with experimental evidence, but not, for example, evolutionary psychology), but there is no real pluralism in the types of psychology that BPP draws upon. The relationship between BPP in practice and other disciplinary approaches is hence what we could describe as one of select interdisciplinarity. This select interdisciplinarity does not acknowledge considerable diversity within the contributing disciplines in advocacy of behavioural applications.

### **From “nudge” to behavioural practice**

The lesson to draw from this is that ethicists of public policy should be aware that behavioural public policy comes with novel claims about what behavioural sciences *are*

and what they are *for*; it also comes with (implicit) claims about the role of science in policy, together with novel institutional configurations (nudge units) and organisational interests that bring these changes about. Therefore, for ethicists to make adjudications on behavioural policy, it is necessary to understand the empirical reality of the practice they want to evaluate. Informed by this approach, we highlight three sets of challenges BPP poses in democracies through the specific (1) *instruments* it uses, the (2) *organisational forms* it involves, and the (3) *values* that scientific behavioural research is permeated with. These facets are not necessary features of BPP but are historically grown and therefore contingent features that characterise the currently dominant variant of bringing “psychological insights to policy”.

The purpose here is hence not to claim irreconcilability of behavioural approaches with democratic principles on a general level, but to investigate some contingent practices insofar as they stand in tension with democratic principles. In the case of behavioural *instruments*, the behaviour of citizens is impacted without policy-makers necessarily relying on reason-giving, and by relying on non-participatory research methods. In the case of behavioural policy *organisations*, policy is made in a setting that is not under direct democratic control, and administrative interventions have thus far preferred administrative interventions over legislative actions. Lastly, regarding *values embedded in scientific research*, BPP’s emphasis on the practical use of behavioural scholarship confronts policy-makers with only a partial and narrow view of how a science of human behaviour may benefit public policy, which might undercut BPP’s ability to provide fora for scrutinising these values in democratic procedures. Such an analysis might also be of constructive value, helping policy-makers to make behavioural practice more transparent, accountable, and attentive to different inter- and intra-disciplinary inputs.

## Instruments

Nudging has become the most prominent example of behavioural public policy as articulated by Thaler and Sunstein in their best-seller *Nudge* (2008). Nudges are gentle non-coercive policy solutions, explicitly justified by the normative framework of libertarian paternalism: they are social interventions that are choice-preserving but welfare enhancing for individual citizens and can be applied in a range of policy fields.

A nudge [. . .] is any aspect of the choice architecture that alters people’s behaviour in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a mere nudge, the intervention must be easy and cheap to avoid. Nudges are not mandates. Putting fruit at eye level counts as a nudge. Banning junk food does not.

(Thaler & Sunstein 2008: 6)

Sunstein is covering many responses to critics of these debates in subsequent works (2015a; 2015b), most recently in his *The Ethics of Influence* (2016). Here, he carefully defends nudges on a variety of ethical grounds (from traditional “perspectives” of autonomy, self-government, dignity), and in particular, welfarist perspectives. Each of these defences has been challenged. Our focus here will lay on instrument-specific issues that are less often discussed. Nudges are difficult to capture with traditional

typologies of policy instruments (Howlett, 1991; Vedung, 1998; Loer (forthcoming)). Some have claimed that especially those nudges which have non-cognitive characteristics are unlike other instruments of command and control, incentives or information and persuasion. Here, some nudges violate a criterion of democratic publicity: how can a polity endorse instruments that “work in the dark”, and which citizens are not aware of (Hausman & Welch, 2010)? More problematically, does a widespread application of behavioural policy not undercut reason-giving in democracies, as behaviour is impacted based on knowledge about behavioural regularities, but not through normative argumentation (Lepoint & Malecka, 2015)? Proponents have attempted to counter these worries by pointing to empirical acceptance of nudges (Reisch et al., 2017) or by saying that public nudge-for-good might be necessary to counterbalance corporate nudges-for-bad (see Schmidt, 2017) or by arguing that some nudges are trivially uncontroversial (e.g., simplified tax forms).

Proponents usually endorse, implicitly or explicitly, a welfarist stance. Correspondingly, nudge units argue they have positive social impact and that they “nudge for good”. We feel, however, that this focus on nudges misses the point: behavioural practice has moved on. When new behavioural units are being set up today, it is not libertarian paternalist principles that are invoked, but rather, it is the discourse of a more scientifically oriented approach to policy that is being used (which means, at times, more “realistic”, “rigorous”, “evidence-based”). Behavioural instruments are now not primarily offered via a normative defence (freedom-of-choice-preserving, non-coercive) or an economical one (cheap) but through a scientific one. Shafir, in the introduction to the landmark *Behavioural Foundations of Public Policy Handbook*, writes that

a rich body of research conducted over the past three to four decades [ . . . ] has changed the way we understand people [ . . . ] our new understanding, this *new view of the human agent*, might help design and implement better public policy.

(Shafir, 2012: 9, our emphasis)

Behavioural reports by nudge units frequently play down the role of nudges as representing only one behavioural instrument among many (Souza-Loureiro, 2014), and that “behavioural science” provides insights beyond the normative frame of libertarian paternalism. Rather than looking at one new instrument (nudge), behavioural policy crucially alters the way that tools are selected compared to alternative policy approaches. Behavioural policy tools differ from traditional policy tools (carrots, sticks, sermons: material incentives, command-and-control measures, persuasive techniques) not because of their inherent characteristics but because of the procedures that have brought them into the policy process. In the case of behavioural instruments, they were chosen as part of an experimental methodology and their presumed effectiveness in changing citizens’ behaviour alone.

For ethicists of public policy, this means entering new territory, as behavioural policy reorders the ways in which tools are being selected. Effectiveness alone matters for tool selection, which means that tool-intrinsic qualities are not considered as relevant by policy-makers anymore (e.g., it makes no normative difference for proponents whether a ban, tax, or psychological cue is getting people to smoke less, where in the past, aspects other than effectiveness played a role in tool selection). However, the focus on



effectiveness simplifies the purpose of such instruments to only one dimension – the impact on citizens' behaviour.

### Organisational forms

Today, there are BPP proponents working on issues from 'behavioral finance, labor contracts, philanthropy, and the analysis of savings and poverty, to eyewitness identification and sentencing decisions, racism, sexism, health behaviours, and voting, health, environment, and nutrition, to dispute resolution, implicit racism, and false convictions' (Shafir 2014: 1). The portfolio of nudge units mirrors this broad range of topics. Take the most prominent example of the Behavioural Insights Team, which to date has published hundreds of behavioural trials in the UK and abroad. These trials are not done in secret: the BIT is extraordinarily transparent in its policy work as it brings norms of academia into policy. This is true for many larger behavioural insights teams, where trials are pre-registered, null findings are reported, and there is a close cooperation with academic institutions as well as in transdisciplinary projects with NGOs or private companies. In addition, BPP engages openly with their own "behavioural failures", as will be pointed out ahead. Yet, other aspects of BPP are considerably less transparent: how, for example, do nudge units choose their areas of engagement? Behavioural policy units in different countries seem to prioritise quite differently here, choosing different organisational forms (centralised, networked; bottom-up or top-down; academic or administration-driven). They range from Danish grassroots organisation INudgeYou to purely governmental units in Germany (Projektgruppe "Wirksam Regieren") or Japan, to policy initiatives at the international, local, or regional level. The most influential team, the UK BIT, was actually semi-privatised and became a so-called social purpose organisation, with several offices around the world. Others went from a higher involvement of private actors in looser networks to a more public status (Nudge France), or went from being federally organised (SBST in the US) to a more decentralised format (behavioural projects at the Department of Defense, as well as on municipal and city level). At the same time, new behavioural consultancies are being created in the private sector (e.g., at Nestle, Deloitte, Ogilvy).

Yet some organisational forms are more problematic than others. How legitimate is it for behavioural public policy to become a private consulting service? Should behavioural policy units be as public as possible in terms of ownership structure? We think that publicly owned nudge units enable them better to fend off exploitative commercial behavioural practices ("counter-nudges"). An important question going forward here is to figure out which organisational form best accommodates a variety of voices (from different scientific fields and beyond) in behavioural policy practice.

This becomes particularly pressing as BPP units develop their own political economy (Schubert 2014a): teams face resistance from within government and from industry and corporate lobby groups, and constantly have to persuasively present their findings in order to gain political credibility. In practice, this has turned behavioural units into political and strategic actors themselves that have to balance accountability and institutional survival. Fox and Silk (2015: 5), for example, argue for the need to 'learn several lessons from the unrivaled success of economists in influencing policy. We highlight three: Communicate simply, field test and quantify results, and occupy

positions of influence'. This shows, in accordance with predictions from organisational theory in political science, that once behavioural institutions exist, they need a reason to sustain themselves. As mentioned earlier, BPP constitutes a transnational movement to renew policy (Bakett et al. 2016; Jones et al. 2014), where this institutionalisation is coordinated and strategic.<sup>5</sup>

BPP requires a high level of expertise. Here, convincing sociological explanations of the importance of individual proponents of BPP have been put forward (John 2014; Strassheim et al. 2015; Strassheim & Korinek 2016).<sup>6</sup> These individuals are often a hybrid of researchers and policy-makers themselves. Maybe, in the case of BPP, we should cease to speak dichotomously about policy-makers, policy-takers, and scientists somewhat on the outside. Cairney (2016) suggests usage of the term "policy community" to denote that decisions about policy are made by those who influence informally, and those who bear formal responsibility. In the case of BPP, what is also novel is the direction in which "behavioural knowledge" is travelling in the policy community. It is not just from established findings into practice, but just as often, knowledge is co-produced not because policy-makers and behavioural scientists are participating but because members of nudge units are frequently researchers and policy-makers at the same time who have acquired "behavioural expertise" (as Strassheim and Korinek call it), as well as strategic knowledge about policy processes.

General questions on the place of experts in democracies have been well rehearsed elsewhere (Fischer 2009). Yet, there is at least one additional challenge when relying on behavioural experts in democracies that is posed by the involvement of highly specialised experts (so-called choice architects or *planners*) in organisations that are capable of translating findings from the *behavioural* sciences to policy. For behavioural insights to be applied, at the beginning, there is the diagnosis of a policy problem. In the case of BPP, this usually involves identifying stable cognitive biases in the target population and counteracting them through behavioural interventions (which could be heuristics-triggering or heuristics-blocking interventions). By doing this, however, experts necessarily have to endorse a view of what the target population would reasonably want if they were free from the cognitive bias that experts have diagnosed.<sup>7</sup> Here, behavioural expert opinion is used to diagnose behavioural biases in the target population and implements remedies without necessarily using deliberative fora of democratic institutions of collective will formation. Instead, nudge units act by attempting to identify the "real" but hidden preferences of citizens when choosing interventions. This is hotly debated in economic methodology (Korinek 2014). But the practical challenge is put by the existence of nudge units directly: how can behavioural experts influence policy without forestalling democratic procedures of collective will formation? There have always been ethical challenges of experts (see Philip Tetlock 2015 critique of political expertise, on epistemic democracy in the 'good judgment' project), and there has always been the technocratic critique of the idea of engineering good societies, but what is new here is the institutionalised form of bringing like-minded experts together in well-structured organisations that reference, and draw authority from, a new influential science of human behaviour. Our hypothesis is that proponents of BPP are not merely recipients of findings from an established body ("the" behavioural sciences) but rather the first actively to bring into existence new research in the behavioural sciences (e.g., through studies set up by nudge units) and therefore both recipients and participants in the scientific endeavour they draw upon, and draw

legitimation from a field that they partially create themselves through the design of experiments, tests of new interventions, and new policy ventures. By invoking (self-created) behavioural science as a justification, practitioners thereby assume political and epistemic authority to be one and the same thing.

With unconventional policy tools at their disposal, systematic preference of administrative discretion rather than legislative action, and without full democratic control and oversight, politicised behavioural insights teams may unduly blur the borders between politics, policy, and science.

### Values in behavioural research

We have seen how the attempt to bring the behavioural sciences to policy finds specific contours as it moves from theory to practice. The idea of using knowledge about the regularities and tendencies of human behaviour to make policies effective (i.e., to bring about desired effects) presumes a rather simplified view of science (in this case, the behavioural sciences) as a repository of facts or findings that are then being applied in the practical context (see [Jolls, Sunstein, Thaler 1998](#); [Shafir 2012](#); [Sunstein 1997](#)). An important insight from the contemporary discussions in the philosophy of science is that political, social, and ethical values cannot be kept separate from the processes of knowledge production, and interfere with what makes (behavioural) knowledge reliable and, as such, suitable as a basis of policy applications ([Douglas 2009](#)). Proponents of BPP have not yet discussed what this could mean for their project.<sup>9</sup>

Values have an indispensable role in scientific research ([Douglas 2009](#); [Longino 1990](#); [Kourany 2010](#); [Avlic & Nelson 2007](#); [Solomon 2001](#)). Once we recognise and acknowledge this role, we quickly see that the relationship between behavioural findings and their practical use in the policy contexts is complex. Values are entering the very process of producing knowledge within the behavioural sciences already; they do not appear only at the stage of *applying* this knowledge outside science – that is, in policy. It is perceived as relatively uncontroversial to state that non-epistemic values, such as normative and emotive commitments that concern moral and social life, can influence the choice of topics and of goals that research is expected to serve. The real challenge arises with the question whether there is any type of influence of these values on the acceptance of hypotheses and theories. As scientists have to decide whether the evidence is sufficient to support a claim/hypothesis, non-epistemic values are a necessary part of hypothesis testing and theory choice – particularly in the behavioural and social sciences. They help in assessing the consequences of making a mistake while making judgements about the evidential support for a hypothesis ([Rudner 1953](#)). Furthermore, in order to assess the evidential warrant of a hypothesis, scientists have to decide what kind of evidence is relevant for the hypothesis – at least in some cases this decision can be value-laden ([Longino 1990](#)). Longino also points out that moral and social values are allowed to enter into decisions concerning the background assumptions of scientific reasoning ([Longino 1990](#)).

Generally, it is believed that the best the scientific community can do is to make values explicit and to work out the procedures and approaches for discussing them ([Longino 1990](#), [Longino 2002](#)). This, of course, is difficult to achieve with BPP as it is not directly a subject of democratic control in its current state. This is complicated through the use of experimental methodology in all nudge units around the world:

choice architects have biases – this much behavioural proponents recognise. But citizens themselves, who are partaking in experiments, cannot be choice architects at the same time.<sup>10</sup>

When confronted with the question of what should follow from the realisation that values infuse scientific practice, one common suggestion is to democratise science. The proposal here is that the public receives the chance to legitimately contest (1) the direction of scientific research effort, (2) the legitimacy and acceptability of expertise, and (3) the institutional structures for science (public assessment of research agendas, of expertise, of science's institutions) (Douglas 2008). Currently, BPP does not offer democratised structures on any of these three aspects.

What does the foregoing discussion mean specifically in the context of BPP? Behavioural scientists, as all scientists, are also making inferences in the face of uncertainty. Therefore, while making judgements about which evidence is sufficient for supporting or rejecting a hypothesis, behavioural scientists rely on values when it comes to reasoning about possible consequences of making false positive or false negative errors.

Furthermore, in a paper that has not thus far attracted attention of the field, Lacey argues that behavioural scientists are making value judgements when deciding 'which strategy to adopt' (Lacey 2003:209). Lacey claims that value judgements have an indispensable role during the 'adoption of strategy' during the research process. Adoption of strategy means to 'constrain the kinds of theories (hypotheses, regularities) that might be entertained in a given domain of inquiry, thus to specify the kinds of possibilities that may be explored in the course of the inquiry' (Lacey (2003): 212). Gaining empirical knowledge depends on the strategy scientists adopt, which means that the type of strategy chosen influences what phenomena scientists gain empirical knowledge of. Lacey argues that the involvement of values during the adoption of strategy is especially important in the behavioural and cognitive sciences.

In the behavioural and cognitive sciences scientists confront the choice of adopting, for instance, behaviourist, cognitivist, or sociobiological strategies. The behaviourist strategy constrains hypotheses to those that concern lawful relations between behaviours and environments, whereas the cognitivist strategy does so to those that concern representations of mental structures and computational accounts of mental processes. Lacey points out that 'radical behaviorist approaches are partly motivated by the value of furthering our capability to exercise control over human behaviour, and some cognitive psychology approaches are motivated partly by highlighting the values of rationality and freedom' (p. 219). He argues for a pluralism of strategies employed within the behavioural sciences and claims that attempts to extrapolate and generalise one strategy, for instance, in a form like 'all behavior is explicable in terms of behaviorist categories', or 'all mental phenomena are computable' mean in fact the endorsement of metaphysical claims sustained by the fruitfulness in guiding research and, thus, by value commitments.<sup>11</sup>

We have hypothesised earlier that while BPP is programmatically interdisciplinary in its references to the behavioural sciences, it does not allow for such a pluralism as Lacey describes. Proponents of BPP claim to rely on a particular body of research within the behavioural science, for instance: cognitive psychology and behavioural economics. With behavioural trials, they themselves contribute to co-produce this field, while not taking into account alternative approaches within the behavioural sciences – for

example, the socio-economic studies and social epidemiology or non-mainstream psychological schools. Favouring one approach over another can be related, as suggested by Lacey, to the practical aim and attempt of controlling behaviour, or steering it into more rational directions.

There are two main lessons to be learnt from understanding the role values play in science, and specifically in the behavioural sciences. First, we see that the knowledge on the basis of which BPP formulates its policy solution can be value-laden in the sense explained and elaborated earlier. Values that enter the behavioural research can influence the ways in which this research is being applied to policy. For instance, we observe that certain approaches in behavioural sciences become more influential than others, which is related to the ways in which epistemic and non-epistemic values and interests (e.g., in controlling human behaviour) are entangled.<sup>12</sup> Second, BPP is an approach to policy that lacks full integration with democratic institutions (parliaments, scientific oversight committees) and, as such, poses challenges, or threats, to the possibility of a democratic control over behavioural policies. This becomes problematic also when we account for the role that values play in the behavioural research. One important way of making these values explicit and subject to critical scrutiny is by exposing scientific research to public criticism through democratic engagement. Scrutinising values embedded in behavioural research can have significant impact on the ways in which this research is applied within BPP; it can also allow for more diversity of approaches informing BPP.

## Conclusion

The contemporary phenomenon of behavioural public policy is an innovative and increasingly influential approach to public policy which raises unique concerns about its compatibility with democratic ideas, processes, and institutions. We survey a selection of such problems by analysing the instruments it uses, the organisational forms it has brought about, and the role that values play in behavioural research. BPP poses different problems in democracies: it is hard to make compatible with deliberative and participatory approaches in implementation, it often relies on instruments that impact the behaviour of citizens without necessarily relying on reason-giving, and it instrumentalises science for practical use which might occasionally undercut its reliability as a legitimate source of knowledge.

## Funding information

The work of Magdalena Małecka is sponsored by Academy of Finland grant no. 308682.

## Notes

## References

Alemanno, A. and Sibony, A.L. eds., 2015. *Nudge and the law: A European perspective*. Oxford: Hart Publishing.

- Benartzi, S., Beshears, J., Milkman, K.L., Sunstein, C.R., Thaler, R.H., Shankar, M., Tucker-Ray, W., Congdon, W.J. and Galing, S., 2017. Should governments invest More in Nudging? *Psychological Science*, pp. 1041–1055.
- Benartzi, S. and Thaler, R.H., 2013. Behavioral economics and the retirement savings crisis. *Science*, 339(6124), pp. 1152–1153.
- Berndt, C., 2015. Behavioural economics, experimentalism and the marketization of development. *Economy and Society*, 44(4), pp. 567–591.
- [BSPA](#), 2017. Behavioral Science & Policy Association Website. #Mission. Available at: [https://behavioralpolicy.org/about/#\\_mission](https://behavioralpolicy.org/about/#_mission) [Accessed August 7, 2017].
- Cairney, P., 2016. *The politics of evidence-based policy making*. Springer.
- Chriss, J.J., 2015. Nudging and social marketing. *Society*, 52(1), pp. 54–61.
- Cohen, S., 2013. Nudging and informed consent. *The American Journal of Bioethics*, 13(6), pp. 3–11.
- Douglas, H., 2009. *Science, policy, and the value-free ideal*. Pittsburgh: University of Pittsburgh.
- Exec. Order No. 13707, Federal Register, Vol. 80, No. 191, 56365–56367 (2015) of Sep. 15, 2015
- [Fawcett](#), S.B., Bernstein, G.S., Czyzewski, M.J., Greene, B.F., Hannah, G.T., Iwata, B.A., Jason, L.A., Mathews, R.M., Morris, E.K., Otis-Wilborn, A. and Seekins, T., 1988. Behavior analysis and public policy. *The Behavior Analyst*, 11(1), pp. 11–25.
- Fischer, F., 2009. *Democracy and expertise: Reorienting policy inquiry*. Oxford: Oxford University Press.
- Fox, C.R. and Sitkin, S.B., 2015. Bridging the divide between behavioral science & policy. *Behavioral Science & Policy*, 1(1), pp. 1–12.
- French, J., 2011. Why nudging is not enough. *Journal of Social Marketing*, 1(2), pp. 154–162.
- Gigerenzer, G., 2015. On the supposed evidence for libertarian paternalism. *Review of Philosophy and Psychology*, 6(3), pp. 361–383.
- Gorin, M., Joffe, S., Dickert, N. and Halpern, S., 2017. Justifying clinical nudges. *Hastings Center Report*, 47(2), pp. 32–38.
- Halpern, D., 2016. *Inside the nudge unit: How small changes can make a big difference*. London: Random House.
- Hansen, P.G., 2016. The definition of nudge and libertarian paternalism: Does the hand fit the glove?. *European Journal of Risk Regulation*, 7(1), pp.155-174.
- Hausman, D.M. and Welch, B., 2010. Debate: To nudge or not to nudge. *Journal of Political Philosophy*, 18(1), pp. 123–136.
- Haynes, L., Goldacre, B. and Torgerson, D., 2012. *Test, learn, adapt: Developing public policy with randomised controlled trials*. Cabinet Office.
- Helbing, D., Frey, B.S., Gigerenzer, G., Hafen, E., Hagner, M., Hofstetter, Y., van den Hoven, J., Zicari, R.V. and Zwitter, A., 2017. Will democracy survive big data and artificial intelligence. *Scientific American*, Feb 25.
- Hertwig, R. and Grüne-Yanoff, T., 2017. Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science*
- Heukelom, F., 2014. *Behavioral economics. A history*. Cambridge University Press.
- [Hirshleifer](#), D. and Teoh, S.H., 2017. How psychological bias shapes accounting and financial regulation. *Behavioural Public Policy*, 1(1), pp. 87–105.



- Hobbs, J., 2017. Nudging charitable giving: The ethics of nudging in international poverty reduction. *Ethics & Global Politics*, 10(1), pp. 37–57.
- Howlett, M., 1991. Policy instruments, policy styles, and policy implementation. National approaches to theories of instrument choice. *Policy Studies Journal* 19(2): 1–21.
- John, P., Cotterill, S., Hahua, L., Richardson, L., Moseley, A., Smith, G., Stoker, G. and Wales, C., 2011. *Nudge, nudge, think, think: Using experiments to change citizens' behaviours*. London: Bloomsbury.
- John, P., 2014. Policy entrepreneurship in UK central government: The behavioural insights team and the use of randomized controlled trials. *Public Policy and Administration*, 29(3), pp. 257–267.
- John, P., 2015. Behavioural Science, randomized evaluations and the transformation of public policy: The case of the UK government (November 2, 2015). Available at: <http://dx.doi.org/10.2139/ssrn.2685049>
- Jolls, C., Sunstein, C.R. and Thaler, R., 1998. A behavioral approach to law and economics. *Stanford Law Review*, 50, p. 1471.
- Jones, R., Pykett, J. and Whitehead, M., 2013. *Changing behaviours: On the rise of the psychological state*. Cheltenham: Northampton.
- Kahneman, D. and Tversky, A., 1979. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, pp. 263–291.
- Kemmerer, A., Möllers, C., Steinbeis, M. and Wagner, G. eds., 2016. *Choice architecture in democracies: Exploring the legitimacy of nudging*. Nomos.
- Kourany, J.A., 2010. *Philosophy of science after feminism*. Oxford University Press.
- Lacey, H., 2003. The behavioral scientist qua scientist makes value judgments. *Behavior and Philosophy*, pp. 209–223.
- Lepenies, R. and Małecka, M., 2015. The institutional consequences of nudging – nudges, politics, and the law. *Review of Philosophy and Psychology*, 6(3), pp. 427–437.
- Lepenies, R. and Małecka, M., forthcoming. Behaviour change: Extralegal, apolitical, scientific? In: Beck, S. and Strassheim, H. (eds), *Handbook of Behavioural Change and Public Policy*. Cheltenham, UK/Northampton, MA: Edward Elgar.
- Loer, K., forthcoming. The Enzymatic Effect of Behavioural Sciences – What About Policy-Maker's Expectations? In: Beck, S. and Strassheim, H. (eds), *Handbook of Behavioural Change and Public Policy*. Cheltenham, UK/Northampton, MA: Edward Elgar.
- Longino, H.E., 1990. *Science as social knowledge: Values and objectivity in scientific inquiry*. Princeton: Princeton University Press.
- Longino, H.E., 2002. *The fate of knowledge*. Princeton University Press.
- Małecka, M. and Nagatsu, M., forthcoming. How behavioural research has informed consumer law: The many faces of behavioural research. In: Purnhagen, K., Sibony, A. and Micklitz, H. (eds), *Research handbook on methods in consumer law*. Edward Elgar Publishing.
- Małecka, M. and Lepenies, R., 2018. Is the behavioral approach a form of scientific imperialism. In Maki, U., Walsh, A. and Fernández Pinto, M. (eds), *Scientific Imperialism: Exploring the Boundaries of Interdisciplinarity*. London: Routledge.

- Ménard, J.F., 2010. A ‘nudge’ for public health ethics: Libertarian paternalism as a framework for ethical analysis of public health interventions? *Public Health Ethics*, 3(3), pp. 229–238.
- Mitchell, G., 2004. Libertarian paternalism is an oxymoron. *Nw. UL Rev.*, 99, p. 1245.
- Oullier, O., 2013. Behavioural insights are vital to policy-making. *Nature*, 501, pp. 463–463.
- Pykett, J., Jones, R. and Whitehead, M., 2016. *Psychological Governance and public policy: Governing the mind, brain and behaviour*. Oxon: Routledge.
- Quigley, M. 2013. Nudging for health: On public policy and designing choice architecture. *Medical Law Review*, 21(4), pp. 588–621.
- Rebonato, R. 2012. *Taking liberties: A critical examination of libertarian paternalism*. Palgrave Macmillan.
- Reddy, S.G., 2012. Randomise this! On poor economics. *Review of Agrarian Studies*, 2(2), pp. 60–73.
- Reisch, L.A., Sunstein, C.R. and Gwozdz, W., 2017. Beyond carrots and sticks: Europeans support health nudges. *Food Policy*, 69, pp. 1–10.
- Rudner, R., 1953. The scientist qua scientist makes value judgments. *Philosophy of Science*, 20(1), pp. 1–6.
- Saghai, Y., 2013. Salvaging the concept of nudge. *Journal of Medical Ethics*, 39(8), pp. 487–493.
- Schmidt, A.T., 2017. The power to nudge. *American Political Science Review*, 111(2), pp. 404–417.
- Schubert, C., 2017b. Green nudges: Do they work? Are they ethical? *Ecological Economics*, 132, pp. 329–342.
- [Schubert, C.](#), 2017a. Exploring the (behavioural) political economy of nudging. *Journal of Institutional Economics*, pp. 1–24.
- [Shafir, E.](#), 2012. *The behavioral foundations of public policy*. Princeton University Press.
- Solomon, M., 2001. *Social empiricism*. Cambridge, MA: The MIT press, p. 186.
- Sousa Lourenço, J., Ciriolo, E., Almeida, S.R. and Troussard, X., 2016. Behavioural insights applied to policy: European report 2016. *Brussels: European Commission*.
- Strassheim, H., Jung, A. and Korinek, R.-L., 2015. Reframing expertise. In *Moments of valuation*. Oxford: Oxford University Press, pp. 249–268.
- Strassheim, H. and Korinek, R. 2016: Cultivating ‘nudge’: Knowing behavioural governance. In: Freeman, Richard and Voss, Jan-Peter. (Hg.), *Knowing governance (Palgrave Studies in Science, Knowledge and Policy)*. Houndmills: Palgrave Macmillan, pp. 107–126.
- Sugden, R., 2017. Do people really want to be nudged towards healthy lifestyles? *International Review of Economics*, 64(2), pp. 113–123.
- Sunstein, C.R., 1997. Behavioral analysis of law. *The University of Chicago Law Review*, 64(4), pp. 1175–1195.
- Sunstein, C.R., 2015b. The ethics of nudging. *Yale Journal on Regulation*, 32, p. 413.
- [Sunstein, C.R.](#), 2015a. Nudges do not undermine human agency. *Journal of consumer policy*, 38(3), pp. 207–210.
- Sunstein, C.R., 2016. *The ethics of influence: Government in the age of behavioral science*. New York: Cambridge University Press.



- [Sunstein](#), C.R., 2017. Nudges that fail. *Behavioural Public Policy*, 1(1), pp. 4–25.
- Tetlock, P.E., 2017. *Expert political judgment: How good is it? How can we know?* Princeton University Press.
- Thaler, R.H., 2015. *Misbehaving: The making of behavioral economics*. New York: WW Norton & Company.
- Thaler, R.H. and Sunstein, C.R., 2003. Libertarian paternalism. *The American Economic Review*, 93(2), pp. 175–179.
- Thaler, R.H. and Sunstein, C.R. 2008. *Nudge: Improving decisions about health, wealth, and happiness*. New Haven: Yale University Press.
- Thompson, D. F., 2005. *Restoring Responsibility: Ethics in Government, Business and Healthcare*. Cambridge: Cambridge University Press.
- Tversky, A. and Kahneman, D., 1975. Judgment under uncertainty: Heuristics and biases. In: *Utility, probability, and human decision making*. Springer Netherlands, pp. 141–162.
- Vallier, K., 2016. On the inevitability of nudging. *Georgetown Journal of Law & Public Policy*, 14, p. 817.
- Vedung, E., 1998. Policy instruments: Typologies and theories. In: *Carrots, sticks, and sermons: Policy instruments and their evaluation*, 5, pp. 21–58.
- [Weber](#), M., 2017 [1949]. *Methodology of social sciences*. Routledge.
- Whitehead, M., Jones, R., Howell, R., Lilley, R. and Pykett, J., 2014. Nudging all over the world: Assessing the global impact of the behavioural sciences on public policy. *Economic Social & Research Council Report*. Swindon, UK.
- Whitman, D.G. and Rizzo, M.J., 2015. The problematic welfare standards of behavioral paternalism. *Review of Philosophy and Psychology*, 6(3), pp. 409–425.
- Wilkinson, T.M., 2013. Nudging and manipulation. *Political Studies*, 61(2), pp. 341–355.
- [World Bank](#), 2015. *World development report 2015: Mind, society and behavior*. Washington, DC: World Bank.
- Wylie, A. and Nelson, L.H., 2007. Coming to terms with the value (s) of science: insights from feminist science scholarship. In: Harold Kincaid, John Dupre and Alison Wylie (eds), *Value-free science? Ideals and illusions*. pp. 58–86
- Yeung, K., 2017. ‘Hypernudge’: Big Data as a mode of regulation by design. *Information, Communication & Society*, 20(1), pp. 118–136.

---

■ We follow here an approach to ethics of public policy which is inductive in the sense that we start from empirical observations about changes in the practices of public administration in democracies. We have explored an institutionalist perspective on behavioural policy elsewhere ([Cepanec & Malacka 2015](#)). Here, our approach is similar to Thompson (2005) in that we propose to avoid excessive individual framings (“How to nudge for good?”) and instead focus on themes of institutional responsibility. Indeed, we think that the value of social institutions cannot be appraised by referencing their contribution to the achievement of outcomes that have a single dimension (here: effectiveness). Without spelling out an account of the intrinsic normativity of (certain) social institutions, we take here the more minimal position of arguing against their instrumentalisation.

- <sup>2</sup> For behavioural public policy, see especially how academics and policy-makers collaborated in the influential policy report endorsing the “Test, Learn, Adapt” framework (Haynes et al. 2014).
- <sup>3</sup> With impressive results: the automatic savings plan “Save More Tomorrow” by Thaler and O'Donoghue (2013) has boosted retirements savings in the US by more than \$7.4 billion annually, according to their own calculations.
- <sup>4</sup> We must leave out here the fascinating discussion about whether behavioural instruments should be used to support democratic processes. Sunstein (2016) discusses examples of how behavioural insights might be used to, for example, encourage electoral turnout.
- <sup>5</sup> For example, the UK Behavioural Insights Team (BIT) has introduced a mnemonic (‘APPLES’) with which it advises international partners on how to best convince policy makers to establish nudge units. APPLES stands for: the necessity of networking within Administration, support from Politics, well-recruited People, close physical Location to sources of power, a culture of scientific Experimentation, proximity to universities and other institutions of Scholarship.
- <sup>6</sup> ‘[A]gency as well as structure plays an important role in the adoption and diffusion of the ideas from the behavioural sciences’ (John 2014). John here notes that critics often assume too linear a view of the policy process, and that behavioural ideas and evidence are ‘more limited and less uniform’ in their use as commonly assumed.
- <sup>7</sup> Sunstein suggests that no ethical problems arise for experts if citizens have clear self-control problems, have voiced their preferences beforehand, or can be shown to be ex-post content with behavioural interventions (2017).
- <sup>8</sup> This relies on a normative view of what these borders ought to be. Our position here is what we would call anti-instrumental institutionalism: we generally believe that societal institutions (whether law, politics, science, or others) have value independently of their capability to achieve a specific outcome (i.e., to be effectively achieving a one-dimensional goal).
- <sup>9</sup> This section draws on ideas developed in [Expenses and Ethics in Nudging](#).
- <sup>10</sup> For an exception, see [John et al. \(2014\)](#) for attempts to complement nudge with participatory “think” approaches.
- <sup>11</sup> See also Johnson and Orr in this volume.
- <sup>12</sup> The interpretation of how disciplinary values might impact policy is contested: [Schubert \(2017a\)](#) finds, for example, a systematic bias within regulatory agencies in favour of scientific information that seems to support extending regulation.